



ISSN: 3022-5388

JKAI website: <https://accesson.kr/jkaia>DOI: <http://dx.doi.org/10.24225/jkaia.2024.2.2.9>

# 기상 요인과 미세먼지 농도 간의 관계 분석 및 예측

## Analysis and Prediction of the Relationship between Meteorological Factors and Fine Dust Concentration

윤예빈 Yebin YOUN<sup>1</sup>, 김경아 Kyung-A KIM<sup>2</sup>

Received: November 13, 2024. Revised: December 02, 2024. Accepted: December 13, 2024.

### Abstract

Urban air quality significantly impacts life and health, with fine dust (PM10) and ultrafine dust (PM2.5) posing serious health risks. This study investigates the seasonal variations in fine dust concentrations based on meteorological data from 2022 and 2023, including temperature, humidity, and precipitation. A random forest regression model was utilized to analyze the relationship between fine dust levels and meteorological factors. The results revealed that fine dust concentrations were highest during spring and winter, while summer exhibited the lowest levels. This seasonal pattern is attributed to increased precipitation and higher temperatures, which help reduce airborne particulate matter. The findings underscore the predictive potential of meteorological data in estimating fine dust concentrations. This research provides a foundation for improving urban air quality management and developing public health strategies to mitigate the adverse effects of air pollution.

**Keywords:** Fine dust, Ultrafine dust, Air pollution, Seasonal variation, Random forest

**Major Classification Code:** Artificial Intelligence, Environmental Science, Data Analytics

## 1. Introduction

도시의 대기 오염은 생활의 질과 건강에 큰 영향을 미치는 중요한 환경 문제로 떠오르고 있다. 특히 미세먼지(PM10)는 코나 기관지에서 걸러지지 않고 몸속에 스며들어 폐까지 침투하여 천식과 폐 질환의 원인이 될 수 있다. 초미세먼지(PM2.5)의 경우 일반 미세먼지(PM10)보다 넓은 표면적을 가지기 때문에 많은 유해 물질에 달라붙으며, 크기가 작아 혈관을 통해 다른 인체 기관으로 이동한다. 때문에, 일반 미세먼지보다 건강에 치명적이다.

미세먼지 농도는 계절에 따라 변하고 있으며, 기온, 상대습도, 강수량과 같은 기상요인에 영향을 받는다. 본 논문은 미세먼지(PM2.5 와 PM10) 농도의 계절적 변화를 분석하고, 주요 기상 요인 간의 상관관계를 통해 기상 요인이 미세먼지 농도에 미치는 영향을 파악할 것이다. 이를 통해 기상 요인에 따른 미세먼지 농도의 예측 가능성을 확인해보고, 미세먼지 문제에 대한 대응 방안을 찾아보고자 한다.

대기오염은 인체 건강에 좋지 않은 영향을 끼치는데 미세먼지는 특히 호흡기와 심혈관계에 치명적이다. 미세먼지의 작은 입자가 폐나 심장에 염증을 일으키고,

1 First Author. Undergraduate, Department of Big Data Medical Convergence, Eulji University, Republic of Korea. [sky020903@naver.com](mailto:sky020903@naver.com)

2 Corresponding Author or Second Author, MIIC (Artificial Intelligence Research Institute), Republic of Korea. [kyungakim@naver.com](mailto:kyungakim@naver.com)

© Copyright: The Author(s)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

길게는 호흡기와 심혈관계의 질환에 걸릴 위험을 높일 수 있다.

## 2. Main Body

본 논문에서는 2022 년과 2023 년의 기상 데이터와 미세먼지 데이터를 사용하여 계절별 미세먼지 농도 변화를 분석했다. 기상 데이터는 평균 기온, 평균 상대 습도, 월합 강수량 등으로 구성되어 있으며 미세먼지 데이터는 PM2.5 와 PM10 의 농도로 이루어져 있다. 2022 년과 2023 년의 데이터를 병합하고 '년-월' 형식의 날짜 정보를 변환하여 데이터 분석이 가능하도록 데이터 전처리를 하였다.

전처리 된 데이터를 이용하여 PM2.5 와 PM10 의 농도, 기온, 상대습도, 강수량 간의 상관관계를 분석하였다. [그림 1]은 기상요인과 미세먼지 농도 간 상관관계를 히트맵으로 나타낸 것이다. PM2.5 와 PM10 의 농도는 0.78 로 높은 양의 상관관계를 나타내고 있다. 이는 미세먼지의 PM2.5 와 PM10 농도가 비슷한 요인에 의해 영향을 받으며, 농도 변화가 비슷하게 나타난다는 것을 의미하고 있다. 미세먼지의 PM2.5 와 PM10 농도는 평균 기온과 월합 강수량과 음의 상관관계가 나타나고 있는 것을 확인할 수 있다. 이는 기온이 높은 여름에 미세먼지 농도가 감소하는 경향이 있음을 의미한다. 또한 강수량이 증가할수록 미세먼지 농도가 낮아지는 경향이 있다는 사실을 알 수 있다. 반면, 습도와는 양의 상관관계를 나타내고 있어 습도가 높은 계절에 미세먼지의 PM2.5 와 PM10 의 농도가 증가할 수 있음을 알 수 있다.

기온이 높은 여름에 미세먼지가 감소하는 이유는 높은 온도와 강수량이 공기 중의 미세먼지를 씻어내고 대기 흐름을 원활하게 하여 미세먼지 농도를 줄여주기 때문이다. 기온과 강수량이 미세먼지에 미치는 영향은 유기적으로 작용한다. 반면, 습도는 미세먼지가 공기중에 떠다니는 시간과 농도를 높여주기 때문에 양의 상관관계가 발생하는 경향이 보인다.

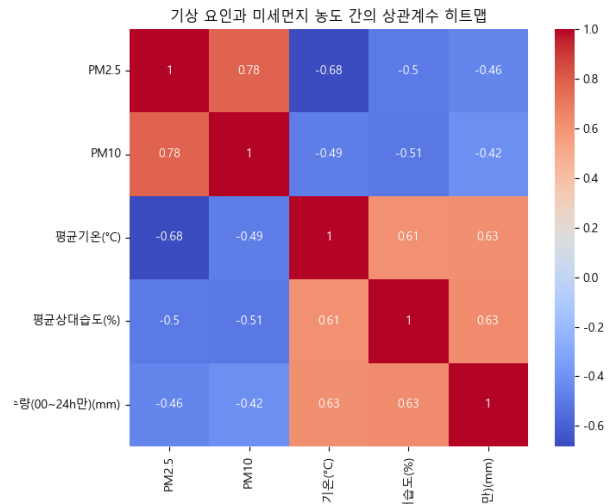


Figure 1: Correlation Heatmap Between Weather Factors and Fine Dust Concentration

미세먼지 농도는 계절에 따라 변화하고 있음을 알 수 있다. 본 논문에서는 봄, 여름, 가을, 겨울의 계절별로 미세먼지의 PM2.5 와 PM10 의 농도 변화를 도시별로 분석하여 미세먼지 농도가 어떤 요인에 의해 변동하는지 확인해보았다.

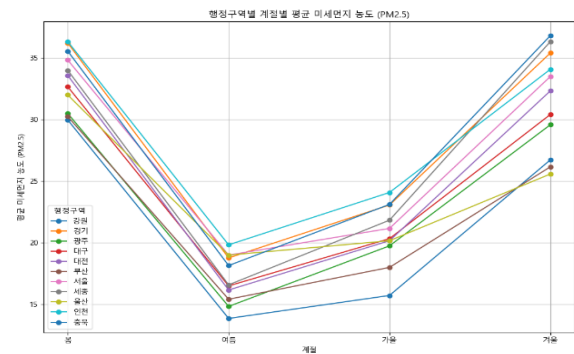


Figure 2: Seasonal Average Fine Dust Concentration (PM2.5) by Administrative Area

[그림 2]를 보면 봄철 PM2.5의 농도가 그래프의 모든 지역에서 높은 것을 확인할 수 있다. 이는 황사와 꽃가루 등과 같은 외부 요인 때문이다. 히트맵 결과를 보다시피 여름철에는 기온 상승과 강수량 증가로 인해 미세먼지 농도가 크게 낮아지는 것을 알 수 있다. 이는 강수로 인해 대기 중의 미세먼지가 씻겨 내려가며, 공기 흐름이 원활

해지기 때문이다. 가을철과 겨울철에는 난방 연료 사용 증가로 인한 대기 오염 때문에 미세먼지 PM2.5의 농도가 높아지는 것을 확인할 수 있다.

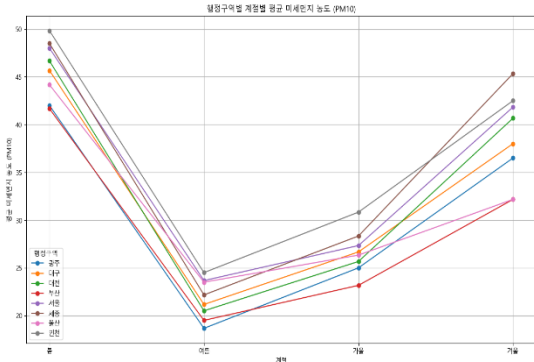


Figure 3: Seasonal Average Fine Dust Concentration (PM10) by Administrative Area

[그림 3]은 PM 10 농도의 계절적 변화를 나타낸다. PM10 은 PM2.5 보다 크기가 큰 미세먼지로, 공업 활동이나 건설 현장에서 발생한다. PM10 농도 역시 PM2.5 와 비슷한 계절적 패턴을 보여주고 있다.

위의 [그림 2]와 [그림 3]은 몇 개의 도시들의 계절별 미세먼지의 평균 농도가 시각화 된 것이다. PM2.5 는 초미세먼지로 크기가 작아서 인체에 미치는 영향이 큰 대기오염 물질이다. PM2.5 농도는 봄과 겨울에 가장 높은 수치를 보였고, 여름에 가장 낮은 농도를 나타내고 있음을 알 수 있다.

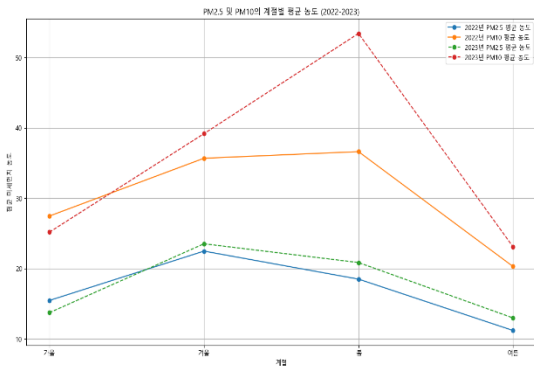


Figure 4: Seasonal Average Concentration of PM2.5 and PM10

위의 [그림 4]는 2022 년과 2023 년의 PM2.5 와 PM10 의 계절별 평균 농도를 비교하여 보여준다.

2022 년과 2023 년 PM2.5 농도: PM2.5 는 여름에 가장 낮고 겨울과 봄에 높은 농도를 보이며 계절에 따라 큰 변동을 보이는 것을 알 수 있다. 특히 2023 년에는 봄에 PM2.5 농도가 상대적으로 더 높게 나타나며, 겨울에서 여름으로 갈수록 PM2.5 농도가 감소하는 패턴이 뚜렷하게 나타난다. 이러한 패턴은 여름철의 강수량과 같은 기상 요인이 미세먼지 농도 감소에 기여할 수 있음을 나타낸다.

2022 년과 2023 년 PM10 농도: PM10 역시 PM2.5 와 유사하게 여름에 가장 낮은 농도를 보이지만, 겨울철과 봄철에 높은 농도를 보이고 있다. 2023 년의 경우 겨울에 PM10 농도가 특히 높게 나타났음을 알 수 있다. 이는 난방과 같은 요인으로 인해 대기 오염이 심화되었을 가능성을 나타낸다.

제시된 모든 그래프에서 봄철 PM 농도가 높은 것을 볼 수 있다. 이는 주로 황사와 꽃가루, 대기 흐름 변화와 같은 외부 요인 때문이다. 황사 발생 빈도와 꽃가루 농도가 높아지면 봄철 미세먼지 농도가 증가한다. 봄철에는 서쪽에서 동쪽으로 부는 바람을 타고 중국와 몽골의 황사가 국내로 들어오면서 미세먼지 농도가 상승하게 된다. 여름철의 미세먼지 농도 감소는 강수량으로 인해 공기에 떠있는 미세먼지가 비에 씻겨 나가고 대기 순환이 원활하게 이루어지기 때문이다.

겨울철 PM 농도가 높아지는 것은 날이 추워짐에 따라 난방을 위한 에너지 소비가 증가하면서 미세먼지 농도가 높아진다. 그 중 석탄과 경유 등의 사용이 증가하여 여기서 미세먼지가 발생하게 된다. 난방에 사용되는 연료에 따른 대기오염은 어떤 연료를 사용하냐에 따라 달라지게되며 PM 농도 차이를 발생시킨다. 겨울철 높은 난방 수요로 인한 미세먼지 농도 증가는 미세먼지 농도의 계절적 패턴에 중요한 원인이 된다.

### 3. Results

본 논문에서는 2022 년과 2023 년의 기상 요인과 미세먼지(PM2.5 와 PM10) 데이터를 바탕으로, 랜덤 포레스트 회귀 모델을 활용하여 미세먼지 농도를 예측했다.

데이터 전처리 및 통합: 2022 년과 2023 년의 기상 데이터를 통합하고, PM2.5 및 PM10 데이터를 장기 형식으로 변환하여 일시와 지역을 기준으로 병합하였다. 랜덤 포레스트 회귀 모델을 사용하여 PM2.5 와 PM10 을 각각 따로 예측하였고, 성능 평가 지표로 평균 제곱 오차(MSE)와 평균 제곱근 오차(RMSE)를 사용했다. [Table 1]은 각각 PM2.5 와 PM10 의 MSE 와 RMSE 의 예측 성능이다.

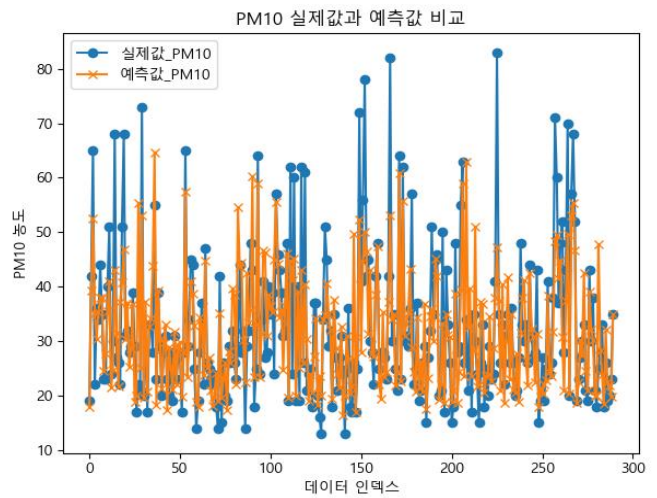
옆의 [Table 1]에서 볼 수 있듯이, 두 가지 미세먼지 농도에 대한 예측 정확성은 비교적 우수하게 나타났다. PM2.5 농도의 경우 오차가 상대적으로 낮아 예측 모델이 미세먼지 농도의 변동을 잘 반영했고 모델의 신뢰도가 높은 것을 알 수 있다. 반면, PM10 농도의 오차는 조금 더 큰 편이지만 여전히 모델의 신뢰도가 높음을 확인할 수 있다.

**Table 1:** MSE and RMSE of PM2.5 and PM10

PM 2.5	
MSE	16.49300
RMSE	4.06115
PM 10	
MSE	79.30451
RMSE	8.90530

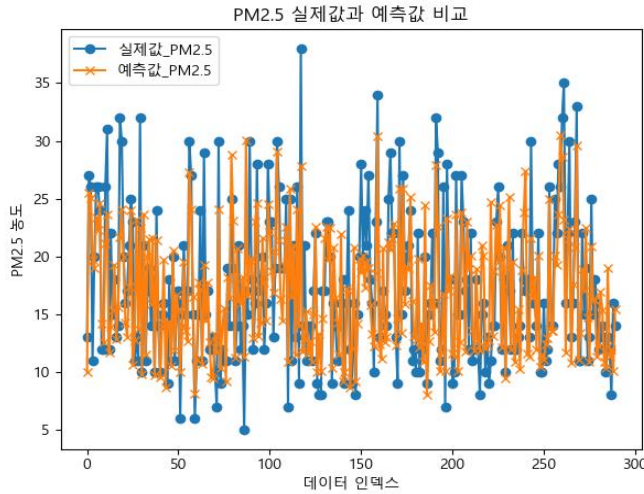
아래의 [그림 5]와 [그림 6]은 PM2.5 와 PM10 농도의 실제 측정값과 랜덤 포레스트 모델에 의해 예측된 값을 비교하여 도시 환경에서 미세먼지 농도를 정확하게 예측할 수 있는지 확인한다. [그림 5]와 [그림 6]은 각 데이터 인덱스에 따라 미세먼지 농도 값을 표시하고, 주황색 선은 예측값을, 파란색 선은 실제 측정값을 나타낸다

옆의 [그림 5]와 [그림 6]은 PM2.5 와 PM10 농도의 실제 측정값과 랜덤 포레스트 모델에 의해 예측된 값을 비교하여 도시 환경에서 미세먼지 농도를 정확하게 예측할 수 있는지 확인한다. [그림 5]와 [그림 6]은 각 데이터 인덱스에 따라 미세먼지 농도 값을 표시하고, 주황색 선은 예측값을, 파란색 선은 실제 측정값을 나타낸다



**Figure 5:** Comparison of Actual and Predicted Values of PM2.5

[그림 5]는 PM2.5 의 실제값과 예측값을 비교한 것이다. 그래프에서 볼 수 있듯이, 예측값은 실제값을 비교적 잘 따라가고 있으며, 두 값 간의 오차가 크게 벌어지지 않는 경향을 보인다. 대부분의 경우 예측값이 실제 값의 변동성을 반영하여 비슷한 범위 내에 위치하고 있다. 특히, PM2.5 의 농도가 낮은 구간에서는 예측값과 실제값의 차이가 거의 없어, 모델이 낮은 농도 구간에서 높은 정확도를 보이고 있음을 나타낸다



**Figure 6:** Comparison of Actual and Predicted Values of PM10

위의 [그림 6]은 PM10의 실제값과 예측값을 비교한 것이다. 이 경우에도 예측값이 실제값을 따라가는 경향을 보이나, PM2.5보다 다소 높은 오차를 보이고 있다. 특히 PM10의 농도가 높은 구간에서는 실제값 간의 차이가 비교적 크게 나타나고 있다. 이는 PM10의 농도 예측이 PM2.5에 비해 상대적으로 어렵다는 점을 보여주고 있으며, 모델 개선 여지가 있음을 나타낸다.

[그림 5]과 [그림 6]은 모델이 기상 요인과 미세먼지 농도 간의 상관관계를 학습하여 예측 값을 산출할 수 있음을 보여준다. 실제값과 예측값의 분포가 유사한 양상을 보이는 것으로 보아, 계절적 요인이나 일시적 기상 변화와 같은 외부 요인에 따른 미세먼지 농도의 변동을 랜덤 포레스트 모델이 반하고 있음을 알 수 있다

#### 4. Conclusion

본 논문에서는 2022년과 2023년의 기상 데이터, 미세먼지 (PM2.5와 PM10)의 데이터를 활용하여 계절별 미세먼지 농도의 변동과 주요 기상 요인 간의 상관관계를 분석하고, 랜덤 포레스트 회귀 모델로 미세먼지 농도를 예측했다.

분석 결과, PM2.5와 PM10의 농도는 계절에 따라 변하며, 기상요인과 밀접한 관련이 있음을 확인할 수 있었다. 특히, 봄과 겨울철에 미세먼지 농도가 높고, 여름철에는 기온 상승과 강수량 증가로 인해 농도가 낮아지는 패턴이 나타났다.

랜덤 포레스트 회귀 모델을 통한 예측 결과, PM2.5는 상대적으로 낮은 오차를 보여 예측 정확도가 높았고, PM10은 약간 더 높은 오차를 보였지만 신뢰할 수 있는 예측 결과를 제시했다. 실제값과 예측값 간의 비교를 통해, 모델이 계절적 요인에 따른 미세먼지 농도 변화를 일정 부분 반영하고 있음을 확인할 수 있었다.

결과적으로 미세먼지 농도 변화가 계절적 기상 요인에 크게 영향을 받으며, 기상 데이터와 미세먼지 농도를 결합한 머신러닝 모델을 통해 예측 가능하다는 점을 보여주고 있다. 이는 향후 도시 환경에서의 대기질 관리와 공공 건강 정책 수립에 필요한 기초자료로 활용될 수 있으며, 계절적 미세먼지 농도 변화를 예측하여 적절한 대응 방안을 마련하는데 기여할 것으로 기대하고 있다. 추후에는 다양한 기상 요인을 추가하거나, 다른 머신러닝 모델을 활용하여 예측 정확도를 더욱 향상시킬 수 있을 것으로 기대하고 있다.

#### References

Kim, D., & Kim, G. (2011). Analysis of factors affecting fine dust concentration in Seoul: Focusing on meteorological factors and traffic volume. *Journal of Korean Society for Atmospheric Environment*, 27(4), 347-359.

Kim, D., et al. (2021). Comparative analysis of factors affecting fine dust concentration before and after COVID-19. *Journal of Korea Safety Management*, 21(6), 395-403.

Kim, G., et al. (2007). A study on the effects of meteorological factors on fine dust concentration. *Research Report of Korea Environment Policy Evaluation Institute*, 1-150.

Kim, G., et al. (2020). Estimation of air quality changes due to climate change and related health impacts. *KEI Research Report*, 1-150.

Korea Environment Policy Evaluation Institute. (2020). Estimation of air quality changes and related health impacts due to climate change. *KEI Research Report*, 1-200.

Lee, D., & Lee, S. (2019). Random forest-based prediction model for ultrafine dust concentration. *Proceedings of the Korean Institute of Information Scientists and Engineers Conference*, 129-130.

Lee, D., & Lee, S. (2020). Hourly ultrafine dust concentration prediction using time-series data and random forest. *Soongsil*

*University Academic Journal*, 45-56.

- Oh, H., et al. (2021). Effects of urban spatial structure and meteorological factors on seasonal high-concentration fine dust. *Journal of Korea Planning Association*, 56(3), 5-18.
- Shin, D., & Yoon, M. (2018). Analysis of the correlation between air quality and meteorological factors using SPSS. *Journal of Korean Society for Information and Communication*, 22(4), 789-795.
- Kim, D., et al. (2021). Comparative analysis of factors affecting fine dust concentration before and after COVID-19. *Journal of Korea Safety Management*, 21(6), 395-403.